

AD-A087 901

STANFORD UNIV CA DEPT OF STATISTICS
COMPLETE CONVERGENCE OF SHORT PATHS AND KARP'S ALGORITHM FOR TH--ETC(U)
MAY 80 J M STEELE
TR-285

F/6 12/1
N00014-76-C-0475
NL

UNCLASSIFIED

1 of 1
AD-A087 901



END
DATE
FILMED
9 80
DTIC

LEVEL *11*

AD A087901

DTIC
ELECTE
AUG 15 1980
S D C

(12)

Complete Convergence of Short Paths and
Karp's Algorithm for the TSP

By

J. Michael Steele

TECHNICAL REPORT NO. 285

May 2, 1980

Prepared under Contract
N00014-76-C-0475 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted
for any Purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

DTIC
ELECTE
AUG 15 1980
S C D

Accession For	RTIS GRAL
DOC TAB	Unannounced
Justification	
By	
Distribution/	
Availability Codes	
Dist	A
Avail and/or special	

1. Introduction.

The main objective of the present note is to solve a problem proposed by Weide (1978) concerning the complete convergence of certain random variables associated with Karp's probabilistic analysis of the traveling salesman problem (Karp (1976), (1977)).

To set the problem precisely let X_i , $1 \leq i < \infty$, be independent random variables uniformly distributed on the unit square $[0,1]^2$, and let T_n denote the length of the shortest path (in the usual Euclidean distance) which connects each element of $\{X_1, X_2, \dots, X_n\}$.

It was proved by Beardwood, Halton, and Hammersley (1959) that

$$\lim_{n \rightarrow \infty} T_n / \sqrt{n} = \beta$$

with probability one for a finite constant β . This fact was central to the motivation behind Karp's algorithm, but as Weide (1978) points out the Karp algorithm actually calls for the following stronger result to be proved here:

Theorem 1. There is a constant β such that for all $\epsilon > 0$, one has

$$\sum_{n=1}^{\infty} P(|T_n / \sqrt{n} - \beta| > \epsilon) < \infty.$$

This type of convergence is usually called complete convergence, and Theorem 1 stands in a similar relation to the Beardwood, Halton, Hammersley

Theorem as the Hsu-Robbins Theorem stands in relation to the strong law of large numbers (Lukacs (1968), Hsu and Robbins (1947)). The "easy-half" of the Borel-Cantelli lemma shows that Theorem 1 implies the Beardwood-Halton-Hammersley Theorem and the "hard-half" of the Borel-Cantelli lemma shows how Theorem 1 is necessary in modeling contexts where problems of increased size are generated independently of previous problems. (For a full discussion of independent versus incrementing models for random problems one should consult Weide (1978)).

The proof of Theorem 1 is given in the next section and depends upon shapening a subadditivity argument which has been useful in more general contexts (Steele (1979)). The third section discusses a generalization of Theorem 1.

2. Proof of Theorem 1.

Let $N = N(t)$ denote a Poisson counting process with constant growth rate 1. Also, for any $A \subset [0,1]^2$ let $T_t(A)$ denote the length of the shortest path through the points $A \cap \{X_1, X_2, \dots, X_{N(t)}\}$. The method of proof rest upon developing recursions for the functions

$$\phi(t) = E T_t([0,1]^2), \quad \psi(t) = (E T_t^2([0,1]^2))^{1/2}, \quad \text{and} \quad V(t) = \psi(t) - \phi^2(t).$$

First one notes the following:

Lemma 1. There is a constant c such that for any set $I \subset [0,1]$ with n elements there is a path of length no greater than $c\sqrt{n}$ through I .

This lemma is easily proved, but for a proof which yields a good value of c one can consult Few (1955).

Now let

$$S = T_t([0, \frac{1}{2}]^2) + T_t([\frac{1}{2}, 1]^2) + T_t([\frac{1}{2}, 1] \times [0, \frac{1}{2}]) + T_t([\frac{1}{2}, 1] \times [0, \frac{1}{2}])$$

and note by elementary geometry that $T_t([0, 1]^2) \leq S + 4$. By well-known properties of the planar Poisson process one has that the four summands of S are independent and identically distributed. By scaling one also notes $E T_t([0, \frac{1}{2}]^2) = \frac{1}{2} \varphi(t/4)$ and $(E T_t^2([0, \frac{1}{2}]^2))^{1/2} = \frac{1}{2} \psi(t/4)$. Thus, taking expectations of $(S+4)^2$ one has

$$(2.1) \quad \psi^2(t) \leq \psi^2(t/4) + 3 \varphi^2(t/4) + 16 \varphi(t/4) + 16.$$

To simplify (2.1) let $\psi_1(t) = \psi(t^2)$, $\varphi_1(t) = \varphi(t^2)$, and $V_1(t) = V(t^2)$ and note that (2.1) implies

$$\psi_1^2(2t) \leq \psi_1^2(t) + 3 \varphi_1^2(t) + 16 \varphi_1(t) + 16$$

which entails

$$V_1(2t) \leq V_1(t) + 4 \varphi_1^2(t) - \varphi_1^2(2t) + 16 \varphi_1(t) + 16$$

and dividing by $(2t)^2$,

$$(2.2) \quad V_1(2t)/(2t)^2 - (1/4)V_1(t)/t^2 \leq \varphi_1^2(t)/t^2 - \varphi_1^2(2t)/(2t)^2 + 4(\varphi_1(t)+1)/t^2.$$

Applying (2.2) successively to $t, 2t, 2^2t, \dots, 2^{M-1}t$ and summing yields

$$\sum_{k=1}^M v_1(2^k t)/(2^k t)^2 - \frac{1}{4} \sum_{k=0}^{M-1} v_1(2^k t)/(2^k t)^2$$

$$\leq \varphi_1^2(t)/t^2 - \varphi_1^2(2^M t)/(2^M t)^2 + 4(\varphi_1(t)+1) \left(\sum_{k=0}^{M-1} 2^{-2k} \right) t^{-2}$$

and consequently

$$(2.3) \quad \sum_{k=1}^{\infty} v_1(2^k t)/(2^k t)^2 \leq \frac{4}{3} (v_1(t)/t^2 + \varphi_1^2(t)/t^2 + 8(\varphi_1(t)+1)/t^2) < \infty.$$

In terms of $V(t)$, (2.3) becomes for $u = t^2$

$$(2.4) \quad \sum_{k=0}^{\infty} V(4^k u)/(4^k u) < \infty.$$

Now by Beardwood-Halton-Hammersley and dominated convergence theorems we have $\varphi(4^k u)/(4^k u)^{1/2} \rightarrow \beta$ as $k \rightarrow \infty$. Hence by (2.4) and Chebychev's inequality,

$$(2.5) \quad \sum_{k=0}^{\infty} P(|T_{4^k u}([0,1]^2)/(4^k u)^{1/2} - \beta| \geq \epsilon) < \infty.$$

Now for all $n \geq 4^p$ there is a $4^{p+1} \geq m \geq 4^p$ such that $4^k m \leq n \leq 4^{k+1} m$ for some $k \geq 0$. Hence by Boole's inequality and the monotonicity of $T_t[0,1]$

$$(2.6) \quad \begin{aligned} \sum_{n=0}^{\infty} P(|T_n([0,1]^2)/n^{1/2} - \beta| \geq \epsilon) \\ \leq \sum_{k=1}^{\infty} \sum_{m=4^p}^{4^{p+1}} P(T_{4^k m} \geq (\beta + \epsilon)(4^k m)^{1/2}) \\ + \sum_{k=1}^{\infty} \sum_{m=4^p}^{4^{p+1}} P(T_{4^k m} \leq (\beta - \epsilon)(4^{k+1} m)^{1/2}) \end{aligned}$$

$$\leq \sum_{k=1}^{\infty} \sum_{m=4^p}^{4^{p+1}} P(T_{4^k(m+1)} / (4^k(m+1))^{1/2} \geq (\beta + \epsilon)(1 + 4^{-p})^{-1/2}) \\ + \sum_{k=1}^{\infty} \sum_{m=4^p}^{4^{p+1}} P(T_{4^k m} / (4^k m)^{1/2} \leq (\beta - \epsilon)(1 + 4^{-p})^{1/2}).$$

By choosing p , one can guarantee that $(1 + 4^{-p})^{-1/2}(\beta + \epsilon) > \beta$ and $(1 + 4^{-p})(\beta - \epsilon) < \beta$ so (2.5) insures that the last two sums in (2.6) converge, so

$$(2.7) \quad \sum_{n=0}^{\infty} P(|T_n([0,1]^2)/n^{1/2} - \beta| \geq \epsilon) < \infty.$$

To obtain Theorem 1 from (2.7) we note that

$$(2.8) \quad P(|T_n([0,1]^2)/n^{1/2} - \beta| \geq \epsilon) = \sum_{m=0}^{\infty} P(|T_m/n^{1/2} - \beta| \geq \epsilon) \frac{n^m e^{-n}}{m!} \\ \geq \sum_{m=n}^{\infty} P(T_m/n^{1/2} \geq \beta + \epsilon) \cdot \frac{n^m e^{-n}}{m!} \\ + \sum_{m=0}^n P(T_m/n^{1/2} \leq \beta - \epsilon) \cdot \frac{n^m e^{-n}}{m!} \\ \geq P(T_n/n^{1/2} \geq \beta + \epsilon) \left(\frac{1}{4}\right) + P(T_n/n^{1/2} \leq \beta - \epsilon) \left(\frac{1}{4}\right).$$

In the last inequality above one uses the fact that $P(T_n \geq \lambda)$ is

monotone increasing, and the fact that $\sum_{m \geq n} \frac{n^m e^{-n}}{m!}$ and

$\sum_{m \leq n} \frac{n^m e^{-n}}{m!}$ each exceed $\frac{1}{4}$. The bounds (2.7) and (2.8) complete the proof of Theorem 1.

3. Further Results.

In the previous section the aim was to give the most direct proof possible of the result conjectured by Weide (1978). The complete convergence proved there can easily be extended to the broader context of subadditive Euclidean functionals. Since the method of proving complete convergence in this context is directly parallel to the preceding argument, it is sufficient to state the more general result.

To introduce Euclidean functionals, let L denote a real value function defined on the finite subsets \mathbb{R}^d , $d \geq 2$. It will be assumed that $L(X_1, X_2, \dots, X_n)$ is measurable whenever X_i , $1 \leq i \leq n$, are measurable, and that the following four assumptions are satisfied:

- A1. $L(\alpha x_1, \alpha x_2, \dots, \alpha x_n) = \alpha L(x_1, x_2, \dots, x_n)$ for $\alpha > 0$.
- A2. $L(x_1 + t, x_2 + t, \dots, x_n + t) = L(x_1, x_2, \dots, x_n)$ for $t \in \mathbb{R}^d$.
- A3. $L(x_1, x_2, \dots, x_n, x_{n+1}) \geq L(x_1, x_2, \dots, x_n)$.
- A4. $\text{Var}(L(X_1, X_2, \dots, X_n)) < \infty$ when X_i are independent and uniform on the unit cube $[0, 1]^d$.

There is one further assumption which is needed and it is the only one which is not trivial to verify in most applications.

Suppose that $\{Q_i: 1 \leq i \leq m^d\}$ is the partition of $[0, 1]^d$ into sub-cubes of edge $1/m$ and let $tQ_i = \{x: x=ty, y \in Q_i\}$. The subadditivity assumption is the following:

- A5. There is a constant $C > 0$ such that for all positive integers m and positive reals t one has

$$L(\{x_1, x_2, \dots, x_n\} \cap [0, t]^d) \leq \sum_{i=1}^{m^d} L(\{x_1, x_2, \dots, x_n\} \cap tQ_i) + Ctm^{d-1}.$$

It is proved in Steele (1979) that if L satisfies the preceding five assumptions and $\{X_i\}$ are independent with a uniform distribution in $[0,1]^d$ then

$$\lim_{n \rightarrow \infty} L(X_1, X_2, \dots, X_n) / n^{(d-1)/d} = \beta$$

with probability one for some constant $0 \leq \beta < \infty$.

One can easily check that the preceding result implies the Beardwood, Halton, Hammersley theorem. The main observation to be made here is that this result also can be sharpened as in Section 2 to obtain the following:

Theorem 2. Suppose L is a functional which satisfies assumptions A1-A5. If $\{X_i: 1 \leq i < \infty\}$ are independent and uniformly distributed in $[0,1]^d$, then there is a constant $0 \leq \beta < \infty$ so that for all $\epsilon > 0$

$$\sum_{n=1}^{\infty} P(|L(x_1, x_2, \dots, x_n) / n^{(d-1)/d} - \beta| \geq \epsilon) < \infty.$$

There are many functionals which arise in the theory of algorithms and which satisfy A1-A5. The easiest examples are those associated with the Steiner tree problem and the rectilinear Steiner tree problem, but in essentially any geometric problem which deals with minimized lengths one can find natural functionals which meet the conditions of Theorem 2.

As a final point one should note that both Beardwood, Halton, and Hammersley (1959) as well as Steele (1979) contain results valid for random variables with non-uniform distribution. The approximation processes used in these papers to extend the uniform case can again be applied here although to do so would require considerable space. Since the algorithmic applications and the program begun by B. W. Weide (1978) are ably served by Theorems 1 and 2, these last extensions have been omitted.

References

- Beardwood, J., Halton, J.H., and Hammersley, J.M. (1959). The shortest path through many points. Proc. Camb. Philo. Soc. 55, 299-327.
- Hsu, P.L. and Robbins, H. (1947). Complete convergence and the law of large numbers. Proc. Nat. Acad. Sciences, 25-31.
- Few, L. (1955). The shortest path and shortest road through n points. Mathematika, 2, 141-144.
- Karp, R.M. (1976). The probabilistic analysis of some combinatorial search algorithms. Algorithms and Complexity: New Directions and Recent Results, J.F. Taub, ed. Academic Press, New York, pp. 1-19.
- Karp, R.M. (1977). Probabilistic analysis of partitioning algorithms for the traveling-salesman problem in the plane. Math. of Op. Res. 2, 209-224.
- Lukacs, E. (1968). Stochastic Convergence. D.C. Heath and Co. Lexington, Mass. 48-49.
- Steele, J.M. (1979). Subadditive Euclidean functionals and non-linear growth in geometric probability, Technical Report, Dept. of Statistics, Stanford University.
- Weide, B.W. (1978). Statistical Methods in Algorithm Design and Analysis, Thesis, Dept. of Computer Science, Carnegie-Mellon University.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TECHNICAL REPORT #285	2. GOVT ACCESSION NO. AD-A087901	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Complete Convergence of Short Paths and Karp's Algorithm for the TSP.		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
7. AUTHOR(s) J. Michael Steele		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0475
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Code 456 Arlington, Virginia 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042,267
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE May 1980
		13. NUMBER OF PAGES 9
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Traveling salesman problem; Complete convergence; subadditive processes; subadditive Euclidean functionals.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) PLEASE SEE REVERSE SIDE.		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 63 IS OBSOLETE

S/N 0102-LF-314-5601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

312580

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT NO. 285

COMPLETE CONVERGENCE OF SHORT PATHS AND
KARP'S ALGORITHM FOR THE TSP

Let X_i , $1 \leq i < \infty$, be uniformly distributed in $[0,1]^2$ and
let T_n be the length of the shortest path connecting $\{X_1, X_2, \dots, X_n\}$.
It is proved that there is a constant $0 < \beta < \infty$ such that for all
 $\epsilon > 0$

$$\sum_{n=1}^{\infty} P(|T_n/\sqrt{n}-\beta| > \epsilon) < \infty.$$

This answers a question posed by B. W. Weide.

S/N 0102- LF- 014- 6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DATE
FILMED
9-8